done—I obviously did not read it. As to the surveyors, they presumably did read all the references—including those in the Proceedings of the Uzbek Academy; but they did not measure up to the task of tracing the 1846 volume of the Proceedings of the Imperial Academy of Saint Petersburg; shame on the New York Public Library!

BENOIT MANDELBROT, *Book Review Editor*
Harvard University, Cambridge 38, Mass.

Coding Theorems of Information Theory. By J. WOLFOWITZ. Ergebnisse der Mathematik und ihrer Grenzgebiete. N.F., Heft 31. Springer-Verlag, Berlin-Göttingen-Heidelberg; Prentice-Hall, Englewood Cliffs, N.J., 1961. ix + 125 pp. $9.35.

Fairly recently, two joint authors made a distinction between "information theory in the strict sense" and "information theory in the wide sense." In the first category they included that body of research which has its direct origins in the 1947–8 paper of C. E. Shannon. I do not recall precisely their definition of the second category; presumably it consisted of the complement of the first in whatever the reader chooses to embrace within the term "information theory," unqualified. The book under review lends itself admirably, and in the reviewer's opinion, commendably to this classification scheme; it is strictly strict sense information theory.

This book has been written (to quote the preface) to provide, for mathematicians of some maturity, an easy introduction to the ideas and principal known theorems of a certain body of coding theory. The first comment to be made is that the word "introduction" is overly modest; a reader who absorbs the full content of each of the 124 pages will be in possession of at least 70% say of the known results on coding theorems. Second, the requirement of mathematical maturity should by no means be misunderstood to mean a very extensive knowlege of a number of mathematical disciplines. A reader capable of working through a rigorous course in advanced calculus, say, will need only a modest knowlege of elementary probability theory in order to read most of the book and comprehend all of its essential ideas. He may, of course, feel that what he is reading is not quite an easy introduction; however, his persistance will quickly bring substantial rewards. Let us now examine the contents in some detail.

Chapter 1 contains a succinct but adequate heuristic introduction to the discrete memoryless channel. Chapter 2 (Combinatorial Preliminaries) discusses the properties of generated sequences, which were introduced by the author several years ago, and which play a leading role in many of the proofs. The results presented are polished versions of basic facts which have been used previously in the author's publications. Standard properties of the entropy function are also presented. Chapter 3 discusses the discrete memoryless channel. The coding theorem and strong converse are proved, as well as a sharper form of the converse for the binary symmetric channel. Also included in this chapter is the finite state channel, but defined somewhat differently from what the reviewer had considered customary. In the author's definition, to each state of the channel there corresponds a channel probability function (cpf) and the state of the system, at any instant, is a deterministic function of the preceding state and the most recently transmitted symbol. The more usual definition,which leads to a channel with memory, is considered in Chapter 6. Chapter 4 is devoted to compound channels. By compound channels is meant channels whose behavior may be governed by many different cpf's. First, systems in which the cpf remains fixed during the transmission of a word are considered, in the cases where neither transmitter nor receiver knows which cpf governs the system, or when one or both do. In all cases, the coding theorem and strong converse is proved. Then, systems in which the cpf is stochastically determined are considered; here, in addition to studying the aforementioned cases, the author also considers the result of the transmitter knowing the cpf just at the moment of transmission, and not in advance. Coding theorems and converses are proven in all cases, the converse being strong in all cases but one, for which the strong converse was obtained (independently by H. Kesten and J. L. C. Kemperman) too late for inclusion. Finally, by proving the weak converse, the use of feedback is shown to effect no increase in the capacity of a discrete memoryless channel; again, the strong converse was obtained (by Kesten and Kemperman) too late for inclusion. We should remark, for those who are in the habit of giving an author's casual remark greater weight than their own considered judgement, that the assignment of credit to Dobrusin is, at this point, more generous than accurate, as the context of Dobrusin's result is substantially different from the author's. Chapter 5 considers the discrete finite-memory channel; the coding theorem and strong converse are proved. Chapter 6 discusses channels "with a past history," i.e., with various types of memory. For channels of this sort, of course, every probabilistic statement concerning the future behavior of the channel must be conditioned by some assumption concerning the past behavior of the channel. After some discussion of this point, several examples are considered, not with an eye towards maximum generality, but rather to illustrate how earlier methods, suitably modified, can be of considerable value in these more involved cases. Let us mention here that we have been informed by the author that he has been able to establish the strong converse for the finite-state channel of Section 6.6. Chapter 7, General Discrete Channels, discusses other general techniques for proving coding theorems; that of the reviewer, as polished and extended by Blackwell, Breiman, and Thomasian as well as the author, and Shannon's random coding method. Fano's weak converse is also presented, and then a brief discussion to clarify the difference between a strong and weak converse. It is at this point that the reviewer disagrees with the author's assertion that one ought not to speak of capacity without having proved a strong converse. If only for the fact that the term "capacity" antedates "strong converse" by about a decade, the preceding dictum would appear arbitrary. But in fact, a coding theorem and corresponding weak converse uniquely define a number which it is reasonable (and traditional) to call capacity. Furthermore, the weak converse says that that technique (block coding) which was successful, below capacity, in reducing the probability of error to any positive level, fails above capacity. The strong converse simply adds that the failure is, in the limit of large block length, as bad as it can be. Granted that it permits the capacity to be defined as a limit rather than as a limit of a limit supremum, the reviewer does not feel that one is not justified in speaking of a capacity in its absence. We should mention that there are examples (due to the author!) where the strong converse does not obtain, although the weak converse does.

Chapter 8 deals with the semicontinuous memoryless channel, proving the coding theorem, the strong converse, and a sharper version of the latter, due in this instance to Kemperman. Chapter 9 considers a fully continuous channel, i.e., one in which there is a continuum both of input and output symbols. The specific channel which is discussed has as input alphabet the unit interval and output alphabet the real line. The noise appears as an additive Gaussian random variable; the time variable is discrete. The coding theorem and strong converse are proven in four cases; no restriction on the input sequences; $\sum_1^n x_i^2 < n$, where $(x_1, \cdots, x_n)$ is a typical input sequence; $n(1 - \epsilon) \leqq \sum_1^n x_i^2 \leqq n$, and finally, $\sum_1^n x_i^2 = n$. The proofs are carried out first by the author's methods, and then by an adaptation of Shannon's techniques. Chapter 10, entitled Mathematical Miscellanea, contains a proof, following Thomasian, of the asymptotic equipartition property, and of the admissability of an ergodic input for a discrete finite-memory channel, although neither is used elsewhere.

A. Feinstein
University of Illinois
Urbana, Illinois